

UNIVERSITY OF
ILLINOIS LIBRARY
AT URBANA-CHAMPAIGN
BOOKSTACKS

THE HECKMAN BINDERY, INC.
North Manchester, Indiana

KRL

H
or
V

H CC 1W 22 BEBR
21 FACULTY
20 WORKING
19 PAPER

H CC 1W 8 1989
7 NO. 1540-1554

H CC 1W 330
B385 "CV">
NO. 1540-1554
COP. 2

H CC 7W
<IMPRINT>
U. of ILL.
LIBRARY
URBANA

BINDING COPY

PERIODICAL ☐ CUSTOM ☐ STANDARD ☐ ECONOMY ☐ THESIS ☐		NO. VOLS. THIS TITLE		LEAD ATTACH.	
BOOK: ☐ CUSTOM ☐ MUSIC ☐ ECONOMY ☐ AUTH. 1ST ☐					
ACCOUNT LIBRARY NEW RUB OR TITLE I.D. SAMPLE		FOIL COLOR MATERIAL			
86672 001		5632 WHI 488			
ACCOUNT NAME		ISSN.			
UNIV OF ILLINOIS					
ACCOUNT INTERNAL I.D.					
B01912400		NOTES		WHEEL	
ID #2		BINDING FREQUENCY		SYS. I.D.	
SIX3		1		3	
COLLATING				39256	
35					
ADDITIONAL INSTRUCTIONS					
Dept=STX3 Lot=#20 Item=142 INM=007					
1CR2ST3CR MARK BY # E4 90					
SEP. SHEETS		PTS. BD. PAPER		TAPE STUBS	
CLOTH EXT		GUM		FILLER	
STUB		LEAF ATTACH			
POCKETS		SPECIAL PREP			
PAPER		BUCK		CLOTH	
INSERT MAT		ACCOUNT LOT NO		JOB NO	
PRODUCT TYPE		#20		PIECE NO	
HEIGHT		GROUP CARD		VOL THIS	
11.2		11.2		142	
COVER SIZE		X			

0012475

330
B385
1547 COPY 2

STX

BEBR
FACULTY WORKING
PAPER NO. 1547

On Estimating Continuous
Time Financial Models

Y. K. Tse

THE LIBRARY OF THE

APR 17 1989

UNIVERSITY OF ILLINOIS
URBANA-CHAMPAIGN



College of Commerce and Business Administration
Bureau of Economic and Business Research
University of Illinois Urbana-Champaign

BEBR

FACULTY WORKING PAPER NO. 89-1547

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

March 1989

On Estimating Continuous Time Financial Models

Y. K. Tse, Visiting Scholar
Department of Economics

This paper was written when the author was visiting the University of Illinois at Urbana-Champaign, whose hospitality is gratefully acknowledged.

Abstract

We consider the problem of maximum likelihood estimation of some diffusion processes commonly used in studies on stock price and interest rate movements. We distinguish between the exact continuous process and its discretized approximation. Closed form solutions for the maximum likelihood estimators of some processes are obtained, which should facilitate the estimation of these processes. Our analysis of the asymptotic distribution of the estimates of some processes, as well as the small sample findings of a Monte Carlo experiment, suggest that the errors due to discretization are not serious. Although we are not claiming this finding can be generalized, the result is reassuring, as frequently researchers have to work with the discretized approximation. Empirical estimates of some stock price and interest rate processes are reported in the final part of the paper.



Digitized by the Internet Archive
in 2011 with funding from
University of Illinois Urbana-Champaign

<http://www.archive.org/details/onestimatingcont1547tsey>

1. Introduction

The construction of many financial models is based on assumptions concerning the stochastic movements of some security prices. An important example is the celebrated Black-Scholes (1973) option pricing formula, which assumes that the price of the underlying asset, usually a stock, follows a geometric Brownian motion. Alternative option pricing formulae are obtained if the stock price is assumed to follow a jump process (Merton (1976)) or a constant elasticity of variance diffusion process (Cox and Ross (1976)). In studies of the term structure of interest rates, various versions of diffusion processes describing the movements of instantaneous interest rates have been proposed. The works of Vasicek (1977), Brennan and Schwartz (1979, 1980, 1982) and Cox, Ingersoll and Ross (1985) are of particular interest.

To evaluate the price of a derivative security, the parameters driving the stochastic process of the underlying asset have to be estimated. When the underlying stochastic process is believed to undergo volatile changes, some researchers prefer to use "implied estimates", which make use of current data only. While this approach has the appeals of requiring a small amount of input data and, in some empirical applications, achieving good results, it lacks a firm statistical basis. Furthermore, the "implied estimates" presume the validity of the option pricing model, and hence cannot be used as diagnostics or selection criteria for competing models. A more traditional solution is to use statistical estimates based on historical

data. In particular, the estimates derived from the maximum likelihood principle have well-known optimal large sample properties (see, e.g., Amemiya (1985, chapter 4)). An added advantage of this approach is that a large battery of diagnostics (such as the likelihood ratio test, Wald test and Lagrange multiplier test) and model selection criteria (such as the Akaike information criterion) can be used for discriminating competing models if so desired. In this paper, we shall be concerned with the problem of estimation of financial models described by diffusion processes in the maximum likelihood framework.

Lo (1988) studied the theory of maximum likelihood estimation of Ito processes. He established a characterization theorem of the exact likelihood function of data that are sampled only at discrete time points. However, instead of using the exact maximum likelihood estimates (MLE), most empirical works in the finance literature involving the estimation of diffusion processes either use some ad hoc procedures or an approximate MLE obtained by discretizing the diffusion process. As pointed out by Lo, the discretized MLE is in general inconsistent. The magnitude of the inconsistency depends on the sampling interval and typically decreases as the sampling interval is shortened.

While it is possible to characterize the exact likelihood function, the existence and the derivation of it are by no means guaranteed. Hence in many models of application, we have to rely on a discretized approximation of the likelihood function. It is thus important to know the consequences of the approximation, especially its effects on the size of the inconsistency. In this paper we obtain

closed form solutions for the MLE of some diffusion processes commonly used in the finance literature. We conduct a Monte Carlo experiment to compare the performances of the discretized and exact MLE for the cases when both can be calculated analytically. This limited study suggests that the discretized MLE gives results comparable to the exact MLE, provided the data are based on a judicious choice of sampling interval supported by appropriate sample size. It is also found that the rates of convergence of the estimated parameters within the same model to their asymptotic distribution can be dramatically different. If estimating volatility is the main concern, our findings favor choosing data with a short sampling interval.

In Section 2, we discuss the exact and the discretized MLE. Section 3 summarizes the diffusion processes we are considering and derives the closed form solutions of the MLE of some of these processes. Section 4 reports the findings of a Monte Carlo experiment. In Section 5, we present some empirical estimates of the stochastic processes describing the S&P 500 index and the yield of the U.S. three-month Treasury bill. Some conclusions and discussions are given in Section 6.

2. Exact and Discretized MLE of Diffusion Processes

Consider a variable X_t that is generated by the following diffusion process

$$dX_t = a(X_t, \theta)dt + b(X_t, \theta)dW_t, \quad (1)$$

where W_t is a Wiener process. The coefficients $a(\cdot)$ and $b(\cdot)$ are assumed to be known functions of X_t and θ , which is a $p \times 1$ vector of unknown parameters.¹ Suppose X_t is sampled at $n + 1$ discrete time points $t_0, t_1, \dots, t_n$, separated by equal sampling interval h such that $t_j = t_0 + jh$ for $j = 0, \dots, n$.² Denoting X_j as X_t observed at $t = t_j$ and $X = (X_0, \dots, X_n)$, the joint density function of X is given by

$$f(X; \theta) = f_0(X_0; \theta) \prod_{j=1}^n f_j(X_j; \theta | X_{j-1}), \quad (2)$$

where $f_0(X_0; \theta)$ is the unconditional density of X_0 and $f_j(X_j; \theta | X_{j-1})$ is the conditional density of X_j given X_{j-1} . The exact MLE is the value of θ , say $\hat{\theta}$, that maximizes $f(X; \theta)$ (or equivalently $\ln f(X; \theta)$).

Thus, the solution of the MLE requires the functional form of $f_j(\cdot)$ for $j = 0, \dots, n$. Lo (1988) provided a theorem that characterizes the densities $f_j(\cdot)$ in terms of the solution of a partial differential equation.³ Although this theorem can be used to check educated guesses for the solutions of the densities, solving these densities are still quite intractable in many cases. A useful result that greatly simplifies the solution applies to the case when X can be transformed into a variable Y such that the coefficients $a(\cdot)$ and $b(\cdot)$ of the diffusion process generating Y do not depend on Y . The reducibility conditions for the existence of such a transformation is given by Schuss (1980).

For models the exact likelihood function of which cannot be obtained, an alternative is to consider a discretized approximation.

Thus, we consider the following process as an approximation to equation (1)⁴

$$X_t - X_{t-1} = a(X_{t-1}, \theta)h + b(X_{t-1}, \theta)\varepsilon_t \quad t=1, \dots, n, \quad (3)$$

where ε_t are independently and identically distributed (IID) normal variables with mean zero and variance h . The discretized log-likelihood function can be written as⁵

$$\begin{aligned} \ln f^*(X; \theta) = & -\frac{n}{2} \ln(2\pi h) - \sum_{t=1}^n \ln(b(X_{t-1}, \theta)) \\ & - \frac{1}{2h} \sum_{t=1}^n \frac{(X_t - X_{t-1} - a(X_{t-1}, \theta)h)^2}{b^2(X_{t-1}, \theta)}. \end{aligned} \quad (4)$$

We define the discretized MLE as the value of θ , say $\tilde{\theta}$, that maximizes $\ln f^*(X; \theta)$. Equation (3) can be easily recognized as a non-linear dynamic single equation model with heteroscedastic errors.

Note that the discretized approximation implies that, conditional on X_{t-1} , X_t is normally distributed. This result is of course not true in general. However, for sufficiently small h we would expect $f^*(\cdot)$ to be a good approximation to $f(\cdot)$. Attempting to improve the approximation and examine the performance of the discretized MLE, Dietrich-Campbell and Schwartz (1986) transformed the process (3) in such a way that $b(\cdot)$ does not depend on X .⁶ Then they considered three estimation procedures by evaluating $a(\cdot)$ at X_{t-1} , X_t as well as the average of X_{t-1} and X_t . However, as X_t is correlated with ε_t the latter two procedures inadvertently introduce a regressor which is not orthogonal to the error. Thus their finding that the discretized MLE

are not stable is actually a consequence of the inappropriate use of X_t as an input to the regressor.

3. Some Asset Price Models and Their Estimation

3.1 Examples

We now consider some examples of diffusion processes. Two types of models will be considered: stock price model and interest rate model. In the rest of the paper, we denote X_t as the stock price and r_t as the interest rate. For stock price movement, we consider the following models:

(1) Geometric Brownian (GB) Motion

$$dX_t = \mu X_t dt + \sigma X_t dW_t \quad (5)$$

(2) Constant Elasticity of Variance (CEV) Process

$$dX_t = \mu X_t dt + \sigma X_t^\beta dW_t. \quad (6)$$

The GB process has been extensively used in the literature. The CEV process was first suggested by Cox and Ross (1976). It has the property that the elasticity of the instantaneous variance is equal to the constant 2β for all X_t . MacBeth and Merville (1980) estimated the CEV model using an ad hoc search method, in which the parameters were not estimated simultaneously (see the discussion by Manaster (1980)). Other attempts were made by Christie (1982) and Marsh and Rosenfeld (1983). In Section 5, we shall report some results on the discretized MLE of the CEV process.

For interest rate movement, we consider the following models:

(1) Ornstein-Uhlenbeck (OU) Process

$$dr_t = \alpha(\mu - r_t)dt + \sigma dW_t \quad (7)$$

(2) Cox-Ingersoll-Ross (CIR) Process

$$dr_t = \alpha(\mu - r_t)dt + \sigma\sqrt{r_t}dW_t \quad (8)$$

(3) Brennan-Schwartz (BS) Process

$$dr_t = \alpha(\mu - r_t)dt + \sigma r_t dW_t. \quad (9)$$

These processes have the property of reverting to the mean, where μ represents the steady-state mean level and α is the speed of adjustment coefficient. The OU process was proposed by Merton (1971) and studied by Vasicek (1977). Sanders and Unal (1988) estimated this model and tested for its stability. The CIR process was proposed by Cox, Ingersoll and Ross (1985). Like the OU process, its exact likelihood function is known. The BS process has been extensively applied in the literature (see, e.g., Brennan and Schwartz (1979, 1980, 1982), Courtadon (1982) and Ogden (1987)). However, its exact likelihood function is unknown and has to be approximated by its discretized version.

Here we make a note about the interpretation of r_t . The OU and CIR processes were originally suggested as models to describe the instantaneous rate of interest, which is an unobservable state variable. Using arbitrage-free arguments, the yields of discount bonds of various time to maturity can be obtained so that these models

have their observable counterparts. However, the processes generating the yields depend on an extra parameter: the market price of risk. Empirical works on the BS process do not make fine distinction whether r_t is observable or is just a state variable. In this and the next section, we assume either r_t is observable or a good proxy for it (such as the yield on treasury bill with short time to maturity) is available. This will enable us to focus on the issue of estimation of these processes. In Section 5, we shall discuss the problem of estimating the processes based on observed yields.

3.2 Maximum Likelihood Estimation

MLE can be obtained by maximizing the log-likelihood function with respect to θ . For some of the examples described above, closed form solutions can be derived. The availability of such solutions is summarized in Table 1. When a model does not have a closed form solution, numerical optimization methods can be used. We now consider each of the examples given in Section 3.1. We denote $\ell(\theta)$ and $\ell^*(\theta)$ as the exact and discretized log-likelihood functions, respectively.

Insert Table 1 Here

It is well-known that for the GB process $\ln(X_t/X_{t-1})$ are IID $N((\mu - (\sigma^2/2))h, \sigma^2 h)$. Thus, the exact log-likelihood function is given by (for convenience we drop the irrelevant constant from all log-likelihoods)

$$\ell(\theta) = -\frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2 h} \sum_{t=1}^n \left(\ln \left(\frac{X_t}{X_{t-1}} \right) - \left(\mu - \frac{\sigma^2}{2} \right) h \right)^2. \quad (10)$$

Solving for $\partial \ell(\theta)/\partial \theta = 0$, we obtain

$$\hat{\sigma}^2 = \frac{1}{nh} \sum_{t=1}^n \hat{\varepsilon}_t^2$$

$$\hat{\mu} = \frac{\hat{\sigma}^2}{2} + \frac{\ln(X_n/X_0)}{nh}$$

where

$$\hat{\varepsilon}_t = \ln\left(\frac{X_t}{X_{t-1}}\right) - \frac{\ln(X_n/X_0)}{n}. \quad (11)$$

The asymptotic variance matrix of $(\hat{\mu}, \hat{\sigma}^2)'$ can be obtained by evaluating the expected value of the Hessian matrix of $\ell(\theta)$.

Straightforward calculations show that

$$\sqrt{n} \begin{pmatrix} \hat{\mu} - \mu \\ \hat{\sigma}^2 - \sigma^2 \end{pmatrix} \xrightarrow{D} N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{\sigma^2}{h} + \frac{\sigma^4}{2} & \sigma^4 \\ \sigma^4 & 2\sigma^4 \end{pmatrix} \right]. \quad (12)$$

The discretized version of the GB process can be written as

$$\frac{X_t - X_{t-1}}{X_{t-1}} = \mu h + \varepsilon_t, \quad (13)$$

where $\varepsilon_t \sim \text{IID } N(0, \sigma^2 h)$. The MLE of μ is just the sample mean of $(X_t - X_{t-1})/X_{t-1}$ divided by h . Thus, we obtain

$$\tilde{\mu} = \frac{1}{nh} \sum_{t=1}^n \left(\frac{X_t}{X_{t-1}} - 1 \right)$$

and

$$\tilde{\sigma}^2 = \frac{1}{nh} \sum_{t=1}^n \left(\frac{X_t}{X_{t-1}} - 1 - \tilde{\mu}h \right)^2. \quad (14)$$

The variances of $\tilde{\mu}$ and $\tilde{\sigma}^2$ are estimated by $\tilde{\sigma}^2/(nh)$ and $2\tilde{\sigma}^4/n$, respectively.⁷ The expected values of $\tilde{\mu}$ and $\tilde{\sigma}^2$, as given by Lo (1988), are

$$E(\tilde{\mu}) = \frac{1}{h} (e^{\mu h} - 1)$$

and

$$E(\tilde{\sigma}^2) = \frac{n-1}{nh} e^{2\mu h} (e^{\sigma^2 h} - 1). \quad (15)$$

Hence dropping terms of $O(h^2)$, the asymptotic biases of $\tilde{\mu}$ and $\tilde{\sigma}^2$ are, respectively, $\mu^2 h/2$ and $h(2\mu\sigma^2 + (\sigma^4/2))$. Furthermore, the (true) asymptotic variance of $\sqrt{n}(\tilde{\mu} - \mu)$ is $((\sigma^2/h) + (\sigma^4/2)) + 2\mu\sigma^2$, which exceeds the asymptotic variance of $\sqrt{n}(\hat{\mu} - \mu)$ by $2\mu\sigma^2$.⁸

Some conclusions can be drawn from the above analysis. First, the results show that the asymptotic biases of $\tilde{\mu}$ and $\tilde{\sigma}^2$ are quite small. For a typical stock with $\mu = 0.15$ and $\sigma^2 = 0.09$, we tabulate the asymptotic biases of $\tilde{\mu}$ and $\tilde{\sigma}^2$ for $h = 1, 4$ and 13 weeks in Table 2.⁹ For weekly data, the asymptotic biases are quite negligible. Even for quarterly data, the asymptotic bias of $\tilde{\mu}$ is less than 0.3 percent, although the asymptotic bias of $\tilde{\sigma}^2$ is more significant, coming close to 9 percent of σ^2 . Second, we note that while a small sampling interval reduces the asymptotic biases, it increases the sampling variance of the MLE of μ . For the same typical stock, we tabulate the asymptotic standard errors of $\hat{\mu}$ and $\hat{\sigma}^2$ in Table 3 for $n = 100$ and 200 . The table shows that when $h = 1$, the true drift parameter is less than one standard error away from zero for n as large as 200 . On the other hand, the volatility parameter can be estimated relatively precisely. These results suggest that if it is the volatility parameter that is

of primary interest, we should use short time intervals to sample more observations within the sampling period.

Insert Tables 2 and 3 Here

For the CEV process, the exact likelihood function was derived by Cox (1975) (see also Marsh and Rosenfeld (1983, p. 638)). But as the likelihood function depends on the modified Bessel function, estimation by exact MLE is computationally very expensive. Thus, we shall only consider the following discretized approximation

$$X_t - X_{t-1} = \mu X_{t-1} h + \varepsilon_t, \quad (16)$$

where ε_t are independently distributed as $N(0, \sigma^2 h X_{t-1}^{2\beta})$. The MLE can be obtained by numerically maximizing the log-likelihood function

$$\ell^*(\theta) = -\frac{n}{2} \ln \sigma^2 - \beta \sum_{t=1}^n \ln X_{t-1} - \frac{1}{2\sigma^2 h} \sum_{t=1}^n \frac{(X_t - X_{t-1} - \mu X_{t-1} h)^2}{X_{t-1}^{2\beta}}. \quad (17)$$

Some empirical estimates for the S&P 500 index will be reported in Section 5.

We now turn to the interest rate models. First, we consider the OU process. From results given in Vasicek (1977), conditional on r_{t-1} , r_t is normally distributed with mean $e^{-\alpha h} r_{t-1} + \mu(1 - e^{-\alpha h})$ and variance $\sigma^2(1 - e^{-2\alpha h})/(2\alpha)$. Thus the log-likelihood is

$$\ell(\theta) = -\frac{n}{2} \ln \left(\frac{\sigma^2(1 - e^{-2\alpha h})}{2\alpha} \right) - \frac{\alpha}{\sigma^2(1 - e^{-2\alpha h})} \sum_{t=1}^n (r_t - \mu - (r_{t-1} - \mu)e^{-\alpha h})^2. \quad (18)$$

Concentrating $\ell(\theta)$ with respect to σ^2 , the MLE of μ and α are obtained by minimizing the concentrated residual sum of squares. Then the MLE

of σ^2 can be solved from its relevant first order condition. The solutions are given by

$$\hat{\alpha} = -\frac{1}{h} \ln A$$

$$\hat{\mu} = (\bar{r}_1 - \bar{r}_0)/(1-A) + \bar{r}_0$$

and

$$\hat{\sigma}^2 = \left(\frac{\sum_{t=1}^n \hat{\varepsilon}_t^2}{n} \right) \left(\frac{2\hat{\alpha}}{1 - e^{-2\hat{\alpha}h}} \right), \quad (19)$$

where

$$\bar{r}_0 = \left(\sum_{t=0}^{n-1} r_t \right) / n$$

$$\bar{r}_1 = \left(\sum_{t=1}^n r_t \right) / n$$

$$A = \left(\sum_{t=1}^n (r_t - \bar{r}_1)(r_{t-1} - \bar{r}_0) \right) / \sum_{t=1}^n (r_{t-1} - \bar{r}_0)^2$$

$$\hat{\varepsilon}_t = r_t - \hat{\mu} - (r_{t-1} - \hat{\mu})e^{-\hat{\alpha}h}. \quad (20)$$

The variances of $\hat{\alpha}$, $\hat{\mu}$ and $\hat{\sigma}^2$ can be estimated using the following formulae, which are obtained by evaluating the Hessian matrix,¹⁰

$$\hat{\text{Var}}(\hat{\alpha}) = \frac{\hat{\sigma}^2(1-e^{-2\hat{\alpha}h})}{2h^2\hat{\alpha}e^{-2\hat{\alpha}h} \sum_{t=1}^n (r_{t-1} - \hat{\mu})^2}$$

$$\hat{\text{Var}}(\hat{\mu}) = \frac{\hat{\sigma}^2(1-e^{-2\hat{\alpha}h})}{2\hat{\alpha}(1-e^{-\hat{\alpha}h})n}$$

and

$$\hat{\text{Var}}(\hat{\sigma}^2) = \frac{2\hat{\sigma}^4}{n}. \quad (21)$$

The discretized OU process is represented by the equation

$$r_t - r_{t-1} = \alpha(\mu - r_{t-1})h + \varepsilon_t, \quad (22)$$

where $\varepsilon_t \sim \text{IID } N(0, \sigma^2 h)$. Straightforward evaluation gives the following results

$$\tilde{\alpha} = \frac{1}{h} (1-A)$$

$$\tilde{\mu} = \hat{\mu}$$

$$\tilde{\sigma}^2 = \frac{1}{nh} \sum_{t=1}^n \varepsilon_t^2 \quad (23)$$

where

$$\tilde{\varepsilon}_t = r_t - r_{t-1} - \tilde{\alpha}(\tilde{\mu} - r_{t-1})h,$$

and $\hat{\mu}$ and A are as defined in equations (19) and (20). The variances of $\tilde{\alpha}$, $\tilde{\mu}$ and $\tilde{\sigma}^2$ are given by the following estimates

$$\begin{aligned}\hat{\text{Var}}(\tilde{\alpha}) &= \frac{\tilde{\sigma}^2}{n \sum_{t=1}^n (r_{t-1} - \tilde{\mu})^2} \\ \hat{\text{Var}}(\tilde{\mu}) &= \frac{\tilde{\sigma}^2}{\tilde{\alpha}^2 h n} \\ \hat{\text{Var}}(\tilde{\sigma}^2) &= \frac{2\tilde{\sigma}^4}{n}.\end{aligned}\tag{24}$$

The above results can be used to facilitate Monte Carlo comparison of the exact and discretized MLE. It is also reassuring to find that the discretized MLE of μ is equal to the exact MLE.

We now consider the CIR process. In this case, the exact likelihood function is known and depends on the modified Bessel function of the first kind with fractional real order (see Cox, Ingersoll and Ross (1985, pp. 391-392)). Although theoretically the exact MLE can be computed using numerical methods, the computation is excessive.¹¹ However, the discretized version of the process can be easily estimated. We express the approximation as

$$r_t - r_{t-1} = \alpha(\mu - r_{t-1})h + \sigma\sqrt{r_{t-1}}\epsilon_t\tag{25}$$

which can be rearranged as

$$\frac{r_t}{\sqrt{r_{t-1}}} = \frac{\alpha\mu h}{\sqrt{r_{t-1}}} + \sqrt{r_{t-1}}(1-\alpha h) + \sigma\epsilon_t,\tag{26}$$

where ϵ_t are IID $N(0, h)$. Defining $y_t = r_t/\sqrt{r_{t-1}}$, $z_{t1} = 1/\sqrt{r_{t-1}}$, $z_{t2} = \sqrt{r_{t-1}}$, $\phi_1 = \alpha\mu h$, $\phi_2 = 1-\alpha h$ and $\phi_3 = \sigma^2 h$, equation (26) can be rewritten as

$$y_t = \phi_1 z_{t1} + \phi_2 z_{t2} + \varepsilon_t \quad (27)$$

where $\varepsilon_t \sim \text{IID } N(0, \phi_3)$. The MLE of the regression parameters of (27) are given by the ordinary least squares (OLS) estimates. Using obvious matrix notations the MLE of ϕ , $\tilde{\phi}$, are given by $(\tilde{\phi}_1, \tilde{\phi}_2)' = (Z'Z)^{-1}Z'y$ and $\tilde{\phi}_3 = y'My/n$ where $M = I - Z(Z'Z)^{-1}Z'$. As θ and ϕ form one to one correspondence, the MLE of θ can be calculated from $\tilde{\phi}$.¹² Thus we have

$$\tilde{\alpha} = \frac{1 - \tilde{\phi}_2}{h}$$

$$\tilde{\mu} = \frac{\tilde{\phi}_1}{\tilde{\alpha}h}$$

and

$$\tilde{\sigma}^2 = \frac{\tilde{\phi}_3}{h}. \quad (28)$$

To compute estimates of the asymptotic variance of $\tilde{\theta}$, we again make use of the OLS results. We define the matrix

$$H = \frac{\partial(\alpha, \mu)}{\partial(\phi_1, \phi_2)} = \begin{bmatrix} 0 & -\frac{1}{h} \\ \frac{1}{1-\phi_2} & \frac{\phi_1}{(1-\phi_2)^2} \end{bmatrix} = \begin{bmatrix} 0 & -\frac{1}{h} \\ \frac{1}{\alpha h} & \frac{\mu}{\alpha h} \end{bmatrix}. \quad (29)$$

The variance of $\tilde{\theta}$ can be estimated by¹³

$$\hat{\text{Var}}((\tilde{\alpha}, \tilde{\mu})') = h\tilde{\sigma}^2 \tilde{H}(Z'Z)^{-1} \tilde{H}'$$

and

$$\hat{\text{Var}}(\tilde{\sigma}^2) = \frac{2\tilde{\sigma}^4}{n}, \quad (30)$$

where $\tilde{H}$ is H evaluated at $(\tilde{\alpha}, \tilde{\mu})$.

Finally, we consider the BS process. Despite its widespread application, the exact likelihood function of the BS process is unknown. To derive the discretized MLE, we apply the method used for the CIR process. Rearranging terms, the discretized BS process can be written as

$$\frac{r_t}{r_{t-1}} = \frac{\phi_1}{r_{t-1}} + \phi_2 + \varepsilon_t, \quad (31)$$

where $\varepsilon_t \sim \text{IID } N(0, \phi_3)$ and ϕ is as defined prior to equation (27).

Now we let $y_t = r_t/r_{t-1}$, $z_{t1} = 1/r_{t-1}$ and $z_{t2} = 1$. As the approximate model is just a simple linear regression, the solution for $\tilde{\phi}$ is particularly simple. Substituting $\tilde{\phi}$ into equation (28), we obtain the discretized MLE $\tilde{\theta}$. With appropriately defined Z , equation (30) provides estimates of the variance of $\tilde{\theta}$.

4. Monte Carlo Results

Optimality properties of the MLE, such as consistency and asymptotic efficiency, are applicable only when the sample size is large. As the rate of convergence to the asymptotic distribution varies according to the underlying model, it is difficult to provide rule of

thumb that is suitable for all models. It is thus important to consider small sample distributions of the MLE in order to obtain some information regarding the sample size needed to justify the application of asymptotic results. This information may also affect the choice of the sampling interval. Furthermore, except for the GB process few results are known about the properties of the discretized MLE when the data are actually generated by the exact diffusion process. In this section we examine these issues using a Monte Carlo experiment.

As the solutions of the discretized and exact MLE for the GB and OU processes are available in closed form, we select these processes as objects of the experiment. The set-up of the parameters is as follows: For the GB process, we consider $\mu = 0.12, 0.18$ and $\sigma^2 = 0.0625$. For the OU process, we consider $\alpha = 0.8, \mu = 0.07$ and $\sigma^2 = 0.001225$. These parameters are chosen to represent likely values of typical stock price and interest rate movements. For both processes, we fix $h = 1, 4, 13$ and $n = 100, 200, 400$. Random samples of observations were generated based on the exact diffusion process. The discretized and exact MLE were calculated for each sample, and this was repeated 1000 times. To conserve space, not all results of the experiment are reported. Selected findings are summarized in Tables 4 and 5, which give the means and standard deviations of the MLE from the Monte Carlo sample.

Insert Tables 4 and 5 Here

From Table 4, we see that $\hat{\mu}$ and $\tilde{\mu}$ are quite similar, except that, as expected, $\tilde{\mu}$ is larger on average and shows slight signs of upward bias for h equal to 4 and 13. The precision of estimates of μ is rather low, and it decreases with h . We note that 100 observations of four-weekly data achieve the same accuracy as 400 observations of weekly data, as far as the estimation of μ is concerned. In contrast, the standard deviations of estimates of σ^2 are quite small and they do not vary with h .¹⁴ As expected, $\tilde{\sigma}^2$ is upward biased when h is large. When $h = 13$, the relative bias is about 10 percent. To examine the convergence to normality, we calculated the nominal 95 percent confidence interval for each Monte Carlo sample, based on the asymptotic normal distribution and estimates of variance. The results (not reported here) show that the asymptotic distribution is a good approximation for $n = 100$.

To investigate the use of daily data, we conducted further experiments. The results show that with 100 observations of daily data, the standard deviations of the estimates of σ^2 are approximately the same as those obtained from 100 observations of weekly data. Thus, for the purpose of obtaining estimate of σ^2 as input to the Black-Scholes formula, daily data are recommended, as long as the problem of bid-ask spread associated with thin trading can be properly controlled. Since only a few months of past data are required, so that nonstationarity of the variance is unlikely to pose any serious difficulty, the MLE is preferable to the implied estimate, which suffers from the problems discussed in the Introduction.

From Table 5, it is clear that the rates of convergence are quite different for estimators of different parameters. Standard deviations of the estimates of α are particularly large, and they increase as h decreases. For $h = 1$ and $n = 1000$, we performed an extra experiment and obtained the mean of $\hat{\alpha}$ as 1.0383 and its standard deviation as 0.3745. Thus in practical situations, it is unlikely that we would be able to obtain precise estimates of α . In contrast, estimates of μ and σ^2 converge quite quickly. For $\tilde{\sigma}^2$, there is a downward bias. With $h = 13$, the relative bias is about 20 percent. As the asymptotic variances of estimates of α and μ depend on α , statistical inference concerning these parameters using the asymptotic approximation should be interpreted with care. However, this caveat does not apply to σ^2 .

5. Empirical Results

In this section, we report estimates of the diffusion processes defined in Section 3.1 with real data.

For the stock price processes, we use the S&P 500 index. Weekly observations of the index on Wednesday (or the next working day if Wednesday is a holiday) are obtained from Standard and Poor's Security Price Index Record. We consider sampling intervals of one week and four weeks. For weekly data, the series goes from January 1980 through December 1987. For four-weekly data, the series goes from January 1973 through December 1987. To study the effects of the October 1987 crash, we also estimate the processes for data terminating in June 1987. Exact and discretized MLE of the GB process, as well as discretized MLE of the CEV process, are reported in Table 6.

Insert Table 6 Here

As expected, the exact and discretized MLE of the GB process are very close. Estimates of μ are very sensitive to the sampling period as well as to the crash. This sensitivity is much reduced for the estimates of σ^2 , which are quite stable in the period 1973 through 1987, only to be disturbed by the crash. Estimates of β in the CEV process depend critically on whether the crash is included in the sample. For all four periods considered, the instantaneous variances implied by the CEV process calculated at a value of the index approximately equal to the mean of the index over the sample period are computed. These implied variances are very close to the estimates of σ^2 of the GB process. For the four-weekly data from January 1973 through December 1987, the estimate of β is not significantly different from one. In this case, the estimate of μ is very close to its counterpart for the GB process. In other cases, in which the GB process is rejected against the more general CEV process (i.e., $\beta \neq 1$), the estimates of μ of the GB process are downward biased. If the crash is not included in the sample, estimates of β are less than 1, which implies the variance of stock return decreases as the stock price increases. This finding for the pre-crash period concurs with that of Christie (1982).

For the interest rate processes, we use data of the yields on U.S. Treasury bills with three months to maturity. These yields are the average issuing rates in weekly bill auctions reported in the Federal Reserve Bulletin. As explained in Section 3.1, we assume either the

diffusion processes drive the yield or the underlying state variable driven by these processes can be approximated adequately by the yield. Our data set consists of weekly observations from January 1980 through June 1988, and four-weekly observations from January 1972 through June 1988. MLE of the OU (exact and discretized), CIR (discretized) and BS (discretized) processes are reported in Table 7.

Insert Table 7 Here

We observe that the exact and discretized MLE of the OU process are very close, as is expected from the analysis and Monte Carlo results of Sections 3 and 4. For all four processes, standard errors of the estimates of α are quite large. Estimates of α are lower for the CIR and BS process as compared with the OU process, although their differences are considerably swamped by the large standard errors. In contrast, estimates of μ are comparable for all processes. For the weekly data, the implied estimates of the instantaneous standard deviation evaluated at the estimated value of μ are 0.0279 and 0.0246 for the CIR and BS processes, respectively. For the four-weekly data, the corresponding figures are 0.0272 and 0.0263. Thus, the volatilities at the steady-state mean interest rate implied by the CIR and BS processes are considerably lower than that of the OU process.

Ogden (1987) estimated the BS process using monthly data for the period from June 1977 through June 1985. Instead of using the exact formulae derived in this paper, Ogden obtained the MLE by numerical optimization. His estimates for α , μ and σ^2 are, respectively,

0.6384, 0.1053 and 0.0830. We estimated these parameters using four-weekly data for the same period to obtain the results: 0.6847, 0.1033 and 0.1119. Except for σ^2 , the two sets of estimates are quite close. It is unclear whether the discrepancy in the estimates of σ^2 is due to the difference in the data or to the estimation method.

Using arbitrage argument, Vasicek (1977) derived the stochastic process for bond prices when the underlying state variable is assumed to follow the OU process. Similar results were obtained by Cox, Ingersoll and Ross (1985) when the underlying state variable follows a CIR process. Apart from the parameters α , μ and σ^2 , the derived bond price processes depend on an extra parameter which is interpreted as the market price of risk. For Vasicek's result, where bond yields with different time to maturity are normally distributed, the parameters of the term structure can be estimated using cross section plus time series data. This is a topic for future research.

6. Concluding Comments

We have considered the maximum likelihood estimation of the exact, as well as discretized approximation, of some diffusion processes commonly used in studies on stock price and interest rate movements. We have derived closed form solutions for some maximum likelihood estimators, which would facilitate the estimation of these processes.

Our analysis of the exact and discretized maximum likelihood estimators of the geometric Brownian process and the Ornstein-Uhlenbeck process shows that for sampling interval of up to four weeks, with sample size typically used in many applied studies, the inconsistency

due to the discretized approximation is unlikely to be serious. We have reported some small sample findings from a Monte Carlo experiment. If our experience of these processes can be generalized, the error due to discretized approximation is not a problem of major concern. We have also obtained some empirical estimates for the stock price and interest rate processes. These results were compared against results by other authors in the literature.

In this paper we have focused on the problem of estimation. Other important aspects are model selection and tests of misspecification. Statistical analysis should be applied to detect misspecification errors and select the best model. This step would help to decide the best option pricing model to use. Problems in this area will be left for future investigation.

Footnotes

¹Throughout this paper, we assume $a(\cdot)$ and $b(\cdot)$ to be time invariant. Time varying models can be constructed by letting $a(\cdot)$ and $b(\cdot)$ be functions of t .

²The condition of equal sampling interval is not necessary for the results in this section. It is only assumed for convenience.

³Lo's theorem actually applies to a more general class of Ito processes that admit a jump component. However, we shall not consider processes with a jump component in this paper.

⁴We have replaced j with the more conventional time index t . Note that the use of t as a suffix now applies to the continuous time model (as in equation (1)) as well as to the discrete time model (as in equation (3)). The usage of t should be clear from the context.

⁵We assume throughout this paper that X_0 is a given constant. In other words, the MLE are derived conditional on X_0 . This assumption should not affect the asymptotic properties of the MLE.

⁶Dietrich-Campbell and Schwartz (1986) actually considered a two-equation model of short and long interest rates. Also, after transformation the model should be, strictly speaking, represented by another variable, say Y .

⁷These estimates are based on the discretized process. Although they will not converge to the true variance in general, they are the estimates one would use if one did not have information about the exact likelihood. This remark also applies to equations (24) and (30).

⁸We did not pursue with the calculation of the asymptotic variance of $\sqrt{n}(\tilde{\sigma}^2 - \sigma^2)$, which is very cumbersome. Our conjecture is that it is not dependent on h . The Monte Carlo experiment in Section 4 seems to support this conjecture.

⁹In this paper, all parameters are measured at annualized rates. One year is approximated by 52 weeks. In contrast to using the month as a basic unit, using the week ensures sampling intervals are regular.

¹⁰These formulae are obtained by assigning value of zero to the off diagonal elements of the Hessian matrix, which are of negligible order. Again, we note that the asymptotic variance of $\hat{\sigma}^2$ does not depend on h . This remark also applies to equation (24).

¹¹Brown and Dybvig (1986) estimated the volatility and the implied long rate of the CIR process using U.S. Treasury security prices. They assumed that the actual bond prices deviate randomly from the equilibrium prices implied by the CIR process, and fitted a non-linear regression with cross-section data. Much of the stochastic structure of the CIR process is not captured in their model.

¹²Note that equation (23) can also be derived using this approach.

¹³The asymptotic covariance of $\tilde{\sigma}^2$ and $(\tilde{\alpha}, \tilde{\mu})'$ is zero.

¹⁴Differences in the standard deviations of estimates of σ^2 and μ help to explain why it is empirically easier to detect shifts in σ^2 , but not in μ (see, e.g., Boness, Chen and Jatusipitak (1974)).

References

- Amemiya, T., 1985, Advanced Econometrics, Harvard University Press: Cambridge, MA.
- Black, F., and M. Scholes, 1973, "The Pricing of Options and Corporate Liabilities," Journal of Political Economy 81, 637-659.
- Boness, A. J., A. Chen, and S. Jatusipitak, 1974, "Investigations of Nonstationarity in Prices," Journal of Business 47, 518-537.
- Brennan, M. J., and E. S. Schwartz, 1979, "A Continuous Time Approach to the Pricing of Bonds," Journal of Banking and Finance 3, 133-155.
- Brennan, M. J., and E. S. Schwartz, 1980, "Conditional Predictions of Bond Prices and Returns," Journal of Finance 35, 405-416.
- Brennan, M. J., and E. S. Schwartz, 1982, "An Equilibrium Model of Bond Pricing and a Test of Market Efficiency," Journal of Financial and Quantitative Analysis 17, 301-329.
- Brown, S. J., and P. H. Dybvig, 1986, "The Empirical Implications of the Cox, Ingersoll, Ross Theory of the Term Structure of Interest Rates," Journal of Finance 41, 617-630.
- Christie, A. A., 1982, "The Stochastic Behavior of Common Stock Variances: Value, Leverage and Interest Rate Effects," Journal of Financial Economics 10, 407-432.
- Courtadon, G., 1982, "The Pricing of Options on Default Free Bonds," Journal of Financial and Quantitative Analysis 17, 75-100.
- Cox, J. C., 1975, "Notes on Option Pricing I: Constant Elasticity of Variance Diffusions," Working Paper, Graduate School of Business, Stanford University.
- Cox, J. C., J. E. Ingersoll and S. A. Ross, 1985, "A Theory of the Term Structure of Interest Rates," Econometrica 53, 385-407.
- Cox, J. C., and S. A. Ross, 1976, "The Valuation of Options for Alternative Stochastic Processes," Journal of Financial Economics 3, 146-166.
- Dietrich-Campbell, B., and E. S. Schwartz, 1986, "Valuing Debt Options: Empirical Evidence," Journal of Financial Economics 16, 321-343.

- Lo, A. W., 1988, "Maximum Likelihood Estimation of Generalized Ito Processes with Discretely Sampled Data," Econometric Theory 4, 231-247.
- MacBeth, J. D., and L. Merville, 1980, "Tests of the Black-Scholes and Cox Call Option Pricing Models," Journal of Finance 35, 285-301.
- Manaster, S., 1980, "Tests of the Black-Scholes and Cox Call Option Valuation Models: Discussion," Journal of Finance 35, 301-303.
- Marsh, T. A., and E. R. Rosenfeld, 1983, "Stochastic Processes for Interest Rates and Equilibrium Bond Prices," Journal of Finance 38, 635-646.
- Merton, R. C., 1971, "Optimum Consumption and Portfolio Rules in a Continuous-Time Model," Journal of Economic Theory 3, 373-413.
- Merton, R. C., 1976, "Option Pricing when Underlying Stock Returns are Discontinuous," Journal of Financial Economics 1, 125-144.
- Ogden, J. P., 1987, "An Analysis of Yield Curve Notes," Journal of Finance 42, 99-110.
- Sanders, A. B., and H. Unal, 1988, "On the Intertemporal Behavior of the Short-Term Rate of Interest," Journal of Financial and Quantitative Analysis 23, 417-423.
- Schuss, Z., 1980, Theory and Applications of Stochastic Differential Equations, John Wiley and Sons: New York.
- Vasicek, O. A., 1977, "An Equilibrium Characterization of the Term Structure," Journal of Financial Economics 5, 177-188.

Table 1

Solutions of MLE for Various Diffusion Processes^a

Diffusion _b Processes	Exact Likelihood Function	<u>Closed Form Solution of MLE</u>	
		Exact	Discretized
GB	A	A	A
CEV	A	NA	NA
OU	A	A	A
CIR	A	NA	A
BS	NA	NA	A

Notes:

^aA means available and NA means not available.

^bSee Section 3.1 for the meanings of the codes.

Table 2

Asymptotic Biases of $\tilde{\mu}$ and $\tilde{\sigma}^2$, $\mu = 0.15$ and $\sigma^2 = 0.09$

h (in weeks)	<u>Asymptotic Biases</u>	
	$\tilde{\mu}$	$\tilde{\sigma}^2$

1	0.000216	0.000597
4	0.000865	0.00239
13	0.00281	0.00776

Table 3

Asymptotic Standard Errors of $\hat{\mu}$ and $\hat{\sigma}^2$, $\mu = 0.15$ and $\sigma^2 = 0.09$

n	h (in weeks)	<u>Asymptotic Standard Errors</u>	
		$\hat{\mu}$	$\hat{\sigma}^2$

100	1	0.2164	0.013
	4	0.1084	0.013
	13	0.0603	0.013

200	1	0.1530	0.009
	4	0.0767	0.009
	13	0.0427	0.009

Table 4

Monte Carlo Estimates of GB Process

$$\mu = 0.18, \sigma^2 = 0.0625$$

h (in weeks)	n	Estimates			
		$\hat{\mu}$ (exact MLE)	$\hat{\sigma}^2$	$\tilde{\mu}$ (discretized MLE)	$\tilde{\sigma}^2$
1	100	0.1718	0.0618	0.1724	0.0622
		0.1800	0.0089	0.1805	0.0090
	400	0.1758	0.0623	0.1762	0.0627
		0.0917	0.0045	0.0920	0.0045
4	100	0.1816	0.0624	0.1832	0.0643
		0.0897	0.0085	0.0910	0.0089
	400	0.1807	0.0620	0.1820	0.0639
		0.0454	0.0043	0.0460	0.0046
13	100	0.1788	0.0618	0.1832	0.0681
		0.0501	0.0088	0.0524	0.0103
	400	0.1803	0.0624	0.1845	0.0689
		0.0246	0.0045	0.0257	0.0053

Notes:

The Monte Carlo sample size is 1000. For each case, the first number refers to the sample means and the second number refers to the sample standard deviation. For example, when $h = 1$ and $n = 100$, the mean and standard deviation of the Monte Carlo sample of 1000 observations of $\hat{\mu}$ are, respectively, 0.1718 and 0.1800.

Table 5

Monte Carlo Estimates of OU Process

$$\alpha = 0.8, \mu = 0.07 \text{ and } \sigma^2 = 0.1225 \times 10^{-2}$$

h (in weeks)	n	Estimates				
		$\hat{\alpha}$	$\hat{\mu}(\tilde{\mu})$	$\hat{\sigma}^2(\times 10^2)$	$\tilde{\alpha}$	$\tilde{\sigma}^2(\times 10^2)$
1	100	3.5791	0.0527	0.1247	3.4029	0.1165
		2.5193	0.4228	0.0175	2.2830	0.0162
	400	1.3990	0.0698	0.1237	1.3755	0.1205
		0.7213	0.0154	0.0086	0.6962	0.0084
4	100	1.4118	0.0700	0.1255	1.3189	0.1128
		0.7504	0.0181	0.0184	0.6506	0.0166
	400	0.9281	0.0698	0.1227	0.8931	0.1143
		0.2694	0.0075	0.0087	0.2486	0.0079
13	100	0.9903	0.0701	0.1256	0.8656	0.0989
		0.3504	0.0086	0.0204	0.2621	0.0145
	400	0.8486	0.0699	0.1233	0.7623	0.1004
		0.1502	0.0042	0.0098	0.1206	0.0072

Notes:

See the note of Table 4. In addition, the columns of figures under $\hat{\sigma}^2$ and $\tilde{\sigma}^2$ have been scaled up by 100.

Table 6

Empirical Estimates of Stock Price Processes^a

Sampling Period and Interval (in weeks)	Number of Obser- vations	Model ^b	Parameters		
			μ	σ^2	β
80/1-87/6, 1	390	GB(E)	0.1536(0.0557)	0.0233(0.0017)	
		GB(D)	0.1539(0.0559)	0.0234(0.0017)	
		CEV(D)	0.1557(0.0556)	0.1228(0.0076)	0.8363(0.0200) 0.0229 ^c
80/1-87/12, 1	417	GB(E)	0.1206(0.0600)	0.0289(0.0020)	
		GB(D)	0.1207(0.0595)	0.0284(0.0020)	
		CEV(D)	0.1280(0.0590)	0.0023(0.0001)	1.2429(0.0192) 0.0279 ^c
73/1-87/6, 4	189	GB(E)	0.0759(0.0406)	0.0240(0.0025)	
		GB(D)	0.0761(0.0406)	0.0239(0.0025)	
		CEV(D)	0.0873(0.0397)	0.5240(0.0520)	0.6776(0.0301) 0.0244 ^c
73/1-87/12, 4	195	GB(E)	0.0622(0.0426)	0.0272(0.0028)	
		GB(D)	0.0622(0.0420)	0.0265(0.0027)	
		CEV(D)	0.0621(0.0436)	0.0157(0.0013)	1.0536(0.0295) 0.0262 ^c

Notes:

^aFigures in parentheses are standard errors.

^bE denotes exact and D denotes discretized.

^cThis figure is the instantaneous variance implied by the CEV process evaluated at the value of the index approximately equal to the mean over the sample period.

Table 7

Empirical Estimates of Interest Rate Processes^a

Sampling Period and Interval (in weeks)	Number of Obser- vations	Model ^b	Parameters		
			α	μ	σ
80/1-88/6, 1	442	OU(E)	0.6340(0.3615)	0.0801(0.0182)	0.0336(0.0011)
		OU(D)	0.6302(0.3571)	0.0801(0.0182)	0.0334(0.0011)
		CIR(D)	0.5400(0.3632)	0.0784(0.0568)	0.0997(0.0034)
		BS(D)	0.5325(0.3673)	0.0784(0.0160)	0.3135(0.0105)
72/1-88/6, 4	214	OU(E)	0.6555(0.2734)	0.0812(0.0118)	0.0313(0.0015)
		OU(D)	0.6392(0.2600)	0.0812(0.0118)	0.0306(0.0015)
		CIR(D)	0.4913(0.2488)	0.0821(0.0408)	0.0949(0.0046)
		BS(D)	0.4155(0.2420)	0.0844(0.0176)	0.3120(0.0151)

Notes:

^a Figures in parentheses are standard errors.

^b E denotes exact and D denotes discretized.

HECKMAN
BINDERY INC.



JUN 95

Bound-To-Please® N. MANCHESTER,
INDIANA 46962

UNIVERSITY OF ILLINOIS-URBANA



3 0112 060295992